

Calculating Demographic and Epidemiological Quantities in Alberta by Geo-Political Areas

Geographic Methodology Series No. 1

Prepared by
Health Surveillance Branch
Alberta Health and Wellness
Edmonton, Alberta

Latest Revision: December 2003



For more information contact:

Health Surveillance
Alberta Health and Wellness
P.O. Box 1360
10025 Jasper Avenue
Edmonton, Alberta
T5J 2P4

Phone: 780-422-4518
Fax: 780-427-1470

Series title: Alberta Health and Wellness Geographic Methodology Series

ISBN: 0-7785-2712-3 (online)

Geographic Methodology Series No. 1

Calculating Demographic and Epidemiological Quantities in Alberta by Geo-Political
Areas

ISBN: 0-7785-2713-1 (pdf)

Note: Users of the information presented in this document in the analysis of health data must insure that products conform to the Alberta *Health Information Act*.

Executive Summary

This document is part of a family of reports that illustrates and documents the geographic methods required to properly analyze health data in Alberta. The descriptions and methods used are consistent across these reports. Together they provide all needed information required to properly understand the spatial component of health data.

This report documents in detail the methodologies employed by Alberta Health and Wellness and those employed by Statistics Canada in the creation of health measures. It also describes procedures used by other agencies.

It was initiated in response to apparent discrepancies discovered in health measures calculated at the Regional Health Authority level independently by Alberta Health and Wellness and Statistics Canada. It includes an analysis of the differences in health measures that these different methodologies have created. Suggestions for changes to the manner in which health measures are calculated in future are presented including:

- Examining the potential to geocode Vital Event records using postal codes.
- Entering agreements with Statistics Canada and other provinces to obtain Vital Events records when these events occur in other provinces in order to include them in the calculation of health measures.
- Investigating the potential for creating intermediate levels of geography for the calculation of health measures.
- Encouraging Alberta Revenue and Statistics Canada to create a consistent Postal code file for Alberta.

Acknowledgments

The following team prepared this report:

Alberta Health and Wellness

Erik Ellehoj
Dr. Donald Schopfloch

Geographic Information System Consultant
Biostatistician

The team acknowledges the contributions of the following reviewers in helping to insure completeness and accuracy in the report:

Larry Svenson
Tom de Beyer

Alberta Health and Wellness
Alberta Health and Wellness

Todd Chorney

Alberta Revenue

Gail Brese

Alberta Government Services

Russell Wilkens
Jason Gilmore
Brenda Wannell

Statistics Canada
Statistics Canada
Statistics Canada

Juanita Hatcher

Alberta Cancer Board

Table of Contents

Executive Summary	3
Acknowledgments.....	4
Table of Contents	5
List of Figures	7
List of Tables	8
Introduction.....	9
I. Geo-Political Boundaries for Health Reporting	10
Regional Health Authorities (RHAs).....	10
Enumeration Areas (EA).....	11
Dissemination Areas (DA).....	12
Census Subdivisions	12
Postal Codes.....	13
Alberta Standard Geographic Codes.....	14
II. Relationships among Geo-Political Boundaries.....	15
Alberta Treasury Postal Code Translation File (PCTF)	15
Statistics Canada Postal Code Conversion File (PCCF).....	16
Other Postal Code Files	17
Inconsistencies in the assignment of Postal Codes to EAs.....	18
Aggregating between EA, CSD, and RHA.....	19
Aggregating ASGCs to RHAs	20
III. Alberta Health and Wellness Calculation Procedures	20
Calculating Populations	20
Counting birth and death cases	22
Calculating morbidity (hospitalization) cases.....	24
Calculating claims cases	24
Comments on calculating rates	24

IV. Statistics Canada Calculation Procedures.....	25
Calculating Populations	25
Calculating birth and death cases.....	26
Calculating morbidity (hospitalization) cases.....	28
Calculating claims cases	28
V. Analysis of differences between calculations	29
Methodology	29
Population Differences.....	30
<i>Currency</i>	30
<i>Mailing vs residential addresses</i>	31
<i>RCMP, military populations, and correctional institutions populations</i>	31
<i>Reserve populations</i>	31
<i>Non-inclusion</i>	32
Vital Event Count Differences.....	32
VI. Potential for numerator-denominator bias	34
Remaining Potential for bias.....	35
VII. Other Agencies.....	36
Alberta Cancer Board	36
Incidence	36
Population	36
Mortality	37
Geographic Assignment.....	37
Alberta Health and Wellness, Finance and Corporate Services	37
Canadian Institute for Health Information (CIHI)	38
VIII. Conclusions and Recommendations	39
Glossary	41
Bibliography	43
References for Data Sources.....	44
Appendix 1: Sample Files.....	45
Appendix 2: Challenges for Small Area Geographic Analysis in Alberta	49

List of Figures

Figure 1: Regional Health Authorities in Alberta, 1998.....	11
Figure 2: Multiple Enumeration Areas and a single Postal Code.....	18
Figure 3: Two postal codes assigned to a single Enumeration Area	19
Figure 4: Procedures used by Alberta Health and Wellness to calculate rates with data obtained from Alberta Vital Statistics.....	23
Figure 5: Procedures used by Statistics Canada to calculate rates with data obtained from Alberta Vital Statistics.....	27
Figure 6: Postal Code Area vs. ASGC Area.....	36
Figure 7 Census Sub-Divisions with postal codes (red diamonds) Source: PCTF	49
Figure 8 Enumeration areas and postal codes (red diamonds) Source: PCTF.....	50
Figure 9: Enumeration Areas and postal code locations (green triangles) Source: Canada Post.....	51
Figure 10: Enumeration Area populations Source: Statistics Canada	52
Figure 11: Postal code populations Source: Alberta Health and Wellness registration file	53

List of Tables

Table 1 Comparison of Postal Code Files.....	17
Table 2: Comparison between Alberta Health and Wellness and Statistics Canada Methodologies.....	29
Table 3: CD-based Population Estimates (Statistics Canada) and Postal Code (AHW) based populations.....	30
Table 4: Potential Sources of Population Differences	32
Table 5: ASGC to RHA discrepancies in the urban-rural division.....	33

Introduction

Most health statistics, including those derived from administrative data, depend upon a *count* of individuals with particular health characteristics or who are subject to particular health events. In general, these counts are restricted to a specific time period, and a specific geo-political area. Frequently, a *rate* or *proportion* is also calculated, and this requires that a count be divided by the number of individuals who were at risk for the health characteristic or health event of interest. This requires an estimate of the population within the same specific geo-political area during the same specific time period as was used to establish the count.

In order for such health measures to be valid and comparable, counting processes should reference the same geo-political areas and the same time periods, and to the extent possible use the same counting procedures. As well, the counting procedures should be consistent for numerators and denominators when rate or proportions are to be calculated.

The temporal requirement can be addressed relatively simply because time is broken up into consistent units, such as days, months, and years. All agencies use these consistent units of temporal measurement, though there are sometimes discrepancies regarding when to conduct counts or measure populations (e.g. the distinction between calendar year and fiscal year). Geography is not as simple to deal with because most agencies develop their own set of reporting units that may not be comparable to units used by other agencies.

Consequently there will occasionally be discrepancies between similar measures reported by different agencies. While many of these discrepancies are minor, some are sufficiently large to have caused concern. A recent example involved discrepancies in life expectancy measures by Regional Health Authority in the AHW report *Life Expectancy as a Health Status Measure* (1998) and the Statistics Canada report *How Healthy are Canadians?* (1999).

This report is an attempt to document the methodologies employed by Alberta Health and Wellness in calculations of health measures within geopolitical boundaries. The report also attempts to document the methods used by Statistics Canada and other agencies for comparative purposes. The report includes an analysis of the differences in methodologies and the differences in health measures that these have created, and presents recommendations for changes to the manner in which health measures are calculated in future.

I. Geo-Political Boundaries for Health Reporting

This section summarizes the geographic reporting units that are relevant to the calculation of health measures in Alberta.

Regional Health Authorities (RHAs)

The province of Alberta was divided into 17 Regional Health Authorities (RHAs) by the provincial government in 1994. These authorities are responsible for the provision of health services to the citizens whose residences are part of the authority. For example, Edmonton residences receive health services from the Capital Health Authority (RHA 10), while nearby residents of Spruce Grove receive health services from the Westview Health Authority (RHA 08). To date, most analysis of health information in the province of Alberta has been performed using RHAs as the basic geographic reporting unit.

The RHA boundaries were generally formed from municipal boundaries and natural features. These boundaries were adjusted in 1996, 1998, and again in 2001. In 2003 the number of RHAs was decreased to 9. RHA boundaries do not exactly match any other geo-political boundaries, so data collected using other smaller areas must be combined to calculate statistics for the RHAs. The 2003 RHA boundaries follow municipal boundaries in most cases, but there are three exceptions. Figure 1 presents a map of the Regional Health Authorities, reflecting the 2003 boundaries.

Figure 1: Regional Health Authorities in Alberta, 2003



Enumeration Areas (EA)

Enumeration Areas (EA) were created by Statistics Canada for the purpose of the Federal Census performed every 5 years. An EA was the area canvassed by a single census taker and consists of approximately 200 homes, though some EAs may report fewer than 50 homes. Enumeration Area boundaries were determined after the census was completed and initial data has been tabulated. All population counts from the Federal Census are available at this level and it was the smallest standard reporting unit used by Statistics Canada. Within larger urban centres a more detailed geographic unit is available from a Street Network File (SNF) that contains single line streets, railways, and water features, along with names and address ranges. This unit, called a block face, refers to one side of a street between two consecutive intersections.

The EA boundaries and the EA centroid geographic coordinates assigned to each are available from Statistics Canada. The EA boundaries were formed using existing boundaries and natural features, whenever possible. Municipal boundaries established by the Provinces and rivers are the primary sources of boundaries. The EA boundaries were updated every five years when each population census was completed, and reflected any changes made by provinces to municipal boundaries in the previous five-year period. A centroid is a single unique combination of latitude and longitude coordinates assigned to describe the location of a small area. In rural areas, EA centroids are generally at the approximate visual centres of the EA (as long as that centre doesn't fall inside a body of water). For those areas covered by the SNF, the EA centroids are based on the block face centroids.

EA centroids and block face centroids are generally used to assign geographic coordinates to postal codes, as explained in a later section. An EA map is also shown in appendix 2.

Dissemination Areas (DA)

In the 2001 census, the Dissemination Area (DA) replaced the EA as the basic unit of data dissemination. It is the smallest standard geographic area of which all census data are disseminated from the 2001 census onward. It was introduced because the EA populations had a great deal of variability and there were many EAs with very small populations. This required selective data suppression in order to maintain confidentiality.

The dissemination area is a small, relatively stable geographic unit composed of one or more blocks. The DAs respect the boundaries of the census subdivisions (CSD) and census tracts (CT). DAs therefore remain stable over time to the extent that CSDs and CTs do. DAs are uniform in terms of population size (targeted at 400 to 700 persons) to avoid data suppression. However, some DAs may have lower (including 0) or higher populations in order to respect the boundaries of the CSDs or CTs.

DA centroids are created in the same manner as EA centroids. DA centroids began to be used for assignments of postal codes to geographic coordinates in 2004.

Census Subdivisions

Census subdivisions (CSDs) are created by Statistics Canada as geographic divisions larger than EAs (DAs) and used for reporting of Census information. Since all census boundaries are coterminous, forming a hierarchy of geographic areas such that larger areas are formed as exact aggregates of smaller areas, all EAs (DAs in the 2001 census) within a CSD are completely contained by the CSD and there are no partial regional assignments.

CSDs are formed using self-governing geographic entities. In Alberta this includes counties, Improvement Districts (IDs), Municipal Districts (MDs), Special Areas (SAs),

towns, cities, villages, and Indian Reserves. They are formed using these existing municipal boundaries and are updated every five years as the municipal boundaries are revised (in an expanding city, for example). The boundaries are fixed on January 1st of the Census year.

Postal Codes

Postal Codes are defined and maintained by Canada Post for the sorting and delivery of mail. Postal Codes are also widely used as a means of geo-coding databases for the purposes of location analysis, demographic analysis, and other types of geographical analyses.

The structure of a Postal Code is fixed. The sequence is always Alphabetical character / Number / Alpha (full space) Number / Alpha / Number. The first three characters of the postal code are known as the Forward Sortation Area (FSA). The first character of a postal code is allocated in alphabetic sequence from East to West across Canada and denotes a province, territory or a major sector found almost entirely within the boundaries of a province. (There are several instances where the first digit from the adjoining province is used for some residents or businesses of the other province. In Alberta this phenomenon is observed in Lloydminster). Rural postal codes can be distinguished from Urban postal codes by character "0" (zero) in the second position. The third character in an urban code identifies a postal station or city post office. In a rural code these characters identify a set of post offices in a geographic area.

The last three characters - known as the Local Delivery Unit (LDU), will guide the mail to a specific location. In an urban area, the LDU refers to a few buildings along a street, or to a group of mailboxes in suburban areas, or to a wide area (spaghetti-like route) served by a rural route service from an urban post office. For a rural destination, the LDU refers to the community to which the mail is to be delivered. Postal communities (defined by Canada Post) usually do not respect municipal boundaries (defined by the provinces).

Note: The letters W and Z are not used as the first letters of postal codes; D, F, I, O, Q and U are never used in Canadian postal codes. If one of these letters appears in your records, it is inaccurate and should be corrected.

In new suburban areas, postal codes may be linked to a community mailbox. These boxes can service both odd and even sides of the same street, or different streets, within a 300 metre radius of the community mailbox. Often, rural postal codes represent the location of the place where the mail is sorted and not the final place of delivery. This can mean that one postal code, with multiple positions, may be used to provide service to several small rural towns or parts thereof. Most rural postal codes provide service to parts of more than one legal municipality.

Boundaries are available for Forward Sortation Areas (FSAs). FSAs are small in urban areas. There are between 20 to 30 FSAs in cities of 750,000 and 2 or 3 in cities between

50,000 and 75,000. FSAs are very large in rural areas and so are of limited use for geographic coding.

Each postal code is defined by a set of postal delivery locations, which are listed as address ranges. Boundaries for individual (6 digit) postal codes are not available (Canada Post does not need them for their own purposes). It is difficult to generate boundaries for the nearly random set of dots indicating the location of persons using a set of post office boxes with the same postal code. The challenge only increases when dealing with the spaghetti-like routes from urban post offices, spilling into the map territory of the adjacent rural FSAs.

Canada Post updates postal codes monthly. This creates challenges in maintaining the implied boundaries (or geographic correspondences) for every postal code. Most of the new postal codes created in Alberta are assigned to expanding urban areas

Alberta Standard Geographic Codes

In Alberta, a unique code is created and maintained by Alberta Government Services and assigned by Alberta Vital Statistics to any community where a birth, death, or marriage took place.

While the term Standard Geographic Code (SGC) is sometimes used for these units, this has caused great confusion because Statistics Canada uses SGC as an alternate term to refer to CSD. These codings are not the same, and the term Alberta Standard Geographic Code (ASGC) will be used here for the coding used by Alberta Vital Statistics. Originally, ASGCs were loosely based on Statistics Canada's CSDs. The first digit of the CSD was removed and an additional digit was appended to distinguish the rural and urban component of a community. (If there is insufficient information to make this determination, then the ASGC appropriate to the urban component is assigned).

In assigning an ASGC code, postal code and street address are disregarded and only community name is used. ASGCs designate towns, villages, cities, and some reserves. Since their creation ASGCs have come to be assigned to areas that do not have a corresponding CSD assignment. These include smaller villages, hamlets, and communities that no longer exist.

In those instances where a community name does not have an ASGC, Alberta Government Services performs a manual process to identify the community. If required, a new ASGC code is created. Approximately 15 new ASGCs are created each month, most of them for historical purposes. For example, several individuals were born in Dogpound (a small community near Madden) but this community no longer exists. An ASGC code is needed in order to record it as a birthplace for all the appropriate individuals as their deaths are recorded.

Individuals who were born or died outside the province are assigned a code according to their circumstances. Unique codes are used to identify each province and country.

Further information about geographic boundaries can be obtained from the respective agencies, and most of this information is available from the web. There is a considerably larger set of available boundaries if the list of agencies is expanded beyond Alberta Health and Wellness and Statistics Canada. These are not discussed here.

II. Relationships among Geo-Political Boundaries

This section summarizes the relationships between the geographic units described above and the principle tools that are used to provide cross-references between them.

Alberta Treasury Postal Code Translation File (PCTF)

A Postal Code file typically contains a listing of all postal codes as well as specific information about each. The Postal Code file generally provides a cross-reference between the postal code and Enumeration Area (and therefore the Census Subdivision can also be derived). The community name, postal code status, etc are also listed for each postal code. The manner in which these are assigned, and the frequency of updates, varies according to the agency that creates the file. The Postal Code file produced by Statistics Canada will be referred to as the Postal Code Conversion File (PCCF), while the Postal Code file produced by Alberta Treasury will be referred to as the Postal Code Translation File (PCTF).

The PCTF produced and updated by Alberta Treasury and used by Alberta Health and Wellness was developed in 1986 in response to speed-of-update problems with the Statistics Canada PCCF (which always reflects postal codes at least 6 months old). The PCTF is updated monthly, based on the listings issued by Canada Post. Retired postal codes are assigned a code to indicate their new status but remain in the list for historical purposes¹.

Recall that boundaries for EAs are available from Statistics Canada, but a full set of boundaries for postal codes are not available from any source. In the PCTF produced by Alberta Treasury, the cross-classification of EA to PC is created by examining the address range for each postal code and manually assigning the single most appropriate Enumeration Area (EA) code to that postal code. Similarly, Alberta Treasury examines every new postal code in order to determine the address range using the Street Network File (SNF) for that new postal code. From this address range, a list of corresponding EAs is generated. The single EA that best represents the address range of the postal code is selected to represent the postal code.

¹ Some previously retired postal codes have been re-instated but usually in the same area. There is no allocation rule to ensure that these assignments occur in the same area. In effect, such a shift took place on November 2000 in a rural area where the shift was from Chard (in RHA 16) to Kehewin (in RHA 12).

It is important to note that the process of assigning postal codes to RHAs in this file operates through the EAs. An EA centroid to RHA overlay is performed using a GIS and will remain current for the inter-census period as long as the RHA boundaries do not change. The EA to RHA lookup files used by Statistics Canada and by Alberta Health and Wellness are identical. Since the geographic coordinates for a postal code are assigned in the PCTF by looking up the postal code's assigned EA centroid coordinates, this process is equivalent to obtaining geographic coordinates from the EA centroids, assigning these to the postal codes and then performing the overlays with the postal code data. *To reiterate, differences that might arise due to the use of different Postal Code files are the result of the postal code to EA assignment process, not the process by which EAs are aggregated to RHAs.*

Note that in 2003, Alberta Treasury will replace EAs with DAs and use Dissemination Area centroids for the assignment of all postal codes. This will occur after all of the locations of the DA centroids have been compared against the address range of each postal code.

Statistics Canada Postal Code Conversion File (PCCF)

The PCCF is produced by Statistics Canada and lists all the EAs within the observed service area of each postal code. In many cases (mostly rural but affecting 28% of the population), this results in several EAs assigned to the same postal code.

In the PCCF+², records with postal codes linked to multiple EAs are assigned proportionately to the known population distribution, using a population-weighted probabilistic technique. A full description of this system and methods used appear in the PCCF+ manual. Approximately 17,500 postal codes (about 2% of the total) in all Canada are assigned probabilistically. Most are rural postal codes (active or retired), urban postal codes for rural and suburban route services, or postal codes for post office boxes in urban post offices. Approximately 28% of the Canadian population uses such postal codes, and all records for such cases are assigned probabilistically by the PCCF+.

In the PCCF, the Single Link Indicator (SLI) indicates the best single link among those multiple matches. (Note that the SLI does not always result in the same Postal Code to EA assignment as is found in the Alberta Treasury PCTF, but could be used in a similar way to establish a 1:1 correspondence between Postal Code and EA).

In the PCCF, the geographic coordinates assigned to Postal Codes are block face centroids wherever possible, and generally in urban areas, and EA (DA) centroids where block faces are not defined. The block faces provide greater geographic resolution than EA (DA) centroids (as are used by Alberta Treasury's PCTF). The higher resolution of

² PCCF+, formerly known as Geocodes/PCCF is a SAS-based application. The Canadian Cancer Registries, Canadian Vital Statistics Registries, Statistics Canada, and many other agencies use it. The program is available to all licensees of the PCCF. Results using the Single Link Indicator (SLI) of the PCCF will be different than those generated by PCCF+ (see table 2 of the PCCF+ User's Guide).

the PCCF does not affect membership in RHAs or smaller units relative to the PCTF. Also, the higher precision of the PCCF does not imply that it possesses higher accuracy.

The Statistics Canada PCCF is updated every six months. At any given time this file usually contains two to three thousand fewer records than the PCCF generated by Alberta Treasury because of this temporal delay.

Other Postal Code Files

There are two other Postal Code files available for the Province of Alberta. Desktop Mapping Technologies (a company that sells geographic and marketing data) has enhanced the Statistics Canada PCCF by improving the geographic location of the postal code centroids. They have calculated a gravitational centre of each postal code population and used this instead of the EA and block face centroid. The number of postal codes listed is smaller than those available from the PCTF.

Canada Post has also made available a postal code coordinate file. In this file, they list the postal code, community name, and geographic coordinates of the post office. Currently, there is no commitment that the file will be kept up-to-date on a regular schedule. There are also very few agencies using this file because it is not designed to meet current business processes of health and population measures.

Alberta Health Surveillance and most other provincial government departments use the PCTF file created by Alberta Treasury. Statistics Canada and other federal departments usually select the PCCF generated by Statistics Canada to ensure national consistency. Marketing agencies, retailers, etc usually select the PCCF created by Desktop Mapping Technologies. Table 1 provides a summary of the postal code files:

Table 1 Comparison of Postal Code Files

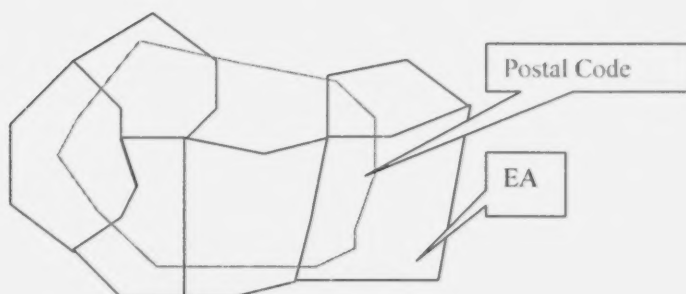
	AHW	Statistics Canada	Desktop Mapping	Canada Post
<i>Elements:</i>	Postal code, EA, community name, activity status, currency period.	Postal code, EA, community name, activity status, currency period, geographic coordinates	Postal code, EA, Geographic coordinates, community name, activity status, currency period	Postal code, Geographic coordinates, community name, activity status, currency period
<i>Developed by:</i>	Alberta Treasury	Statistics Canada	Desktop Mapping/ Statistics Canada	Canada Post

<i>Maintained by:</i>	Alberta Treasury	Statistics Canada	Desktop Mapping	Canada Post
<i>Updates:</i>	Monthly	6 months	6 months	Unknown
<i>Special Notes:</i>	Largest number of postal codes listed. A single EA assigned to each postal code.	Block faces (urban areas). PCCF+ uses a probabilistic method for rural postal codes and for rural services from urban post offices (28% of the population).	Gravitational model to assign centroid location. Lists all block faces or EAs corresponding to the postal code.	Geographic location of post office

Inconsistencies in the assignment of Postal Codes to EAs

In the PCTF, each postal code is assigned to only one single EA, although a single postal code may provide postal service to a region defined by several EAs. For example, a rural postal community may consist of two or three urban EAs and two or more rural EAs, but only one postal code may serve the entire region. Not all EAs have a postal code assigned to them but the served area for the postal code may include several EAs. Figure 2 illustrates this scenario.

Figure 3: Multiple Enumeration Areas and a single Postal Code



In figure 2, a postal code provides service to a region that has been assigned several Enumeration Areas. The EA that is most representative of the postal code region is chosen as the assigned EA in the PCTF. Some of these EAs may not be assigned a postal code under this scenario. The PCCF lists all 7 EAs that overlap with the postal code service area in this rural example and also the best single link or 1:1 match between postal code and EA or blockface. The Postal Code Population Weight File (which is integrated into PCCF+) shows the percentage of the EA population served by each postal code.

There are a few instances where several postal codes are assigned to a single EA. Small rural communities are usually assigned unique postal codes, but when their populations are very small, a single EA may represent two or more villages and surrounding rural area. If each of these two villages has a unique postal code, then both postal codes will be assigned to the same EA. This issue first became apparent in "Assessment of Respiratory Disorders in Relation to Solution Gas Flaring Activities in Alberta" published by Alberta Health Surveillance (1998) where mapping at the postal code level was desired, yet not possible because of this issue. A special file was created for that project which used the geographic coordinates of the postal code community rather than the coordinates derived from the EA centroids.

Figure 4: Two postal codes assigned to a single Enumeration Area

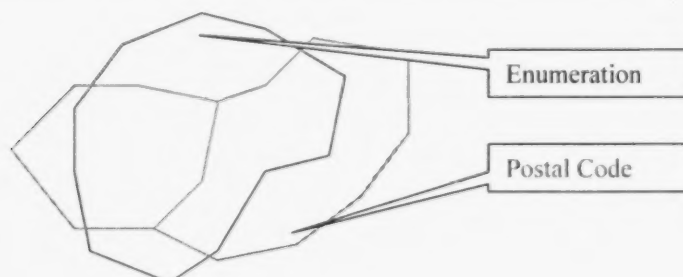


Figure 3 illustrates the issue: a larger EA (in black) contains the majority of the population for two separate postal codes and therefore the assigned coordinates for the two postal codes are identical.

Aggregating between EA, CSD, and RHA

As previously noted, CSDs are formed as aggregates of EAs and in general the EA to CSD cross-references are contained in Postal Code files.

Also as previously noted, the EA to RHA cross-reference table used by Statistics Canada and by Alberta Health and Wellness is identical. In the PCTF, the geographic coordinates for postal codes are the same as those for EAs. The EA to RHA cross-reference table was formed by an overlay of EA to RHA using a GIS and will remain current for the inter-

census period as long as the RHA boundaries do not change. In the 2001 census, the DAs can be aggregated to the CSD level without discrepancies. The PCTF released in 2004 contains DA and EA assignments for every postal code.

Aggregating ASGCs to RHAs

As previously noted, ASGC codes are assigned to every birth and death in Alberta. Address information is collected by individual registration agencies and then sent to Alberta Government Services who assigns an ASGC to each registration record and also creates a cross-reference file. Alberta Treasury also checks this file for accuracy.

In assigning an ASGC code, postal code and street address are disregarded and only community name is used. For each ASGC the cross-reference file lists community name, code, activity status, corresponding CSD and RHA, and other relevant information. The ASGC cases can therefore be aggregated to the RHA level using the ASGC to RHA lookup file. A sample of this file appears in appendix 1.

Please note that aggregating Vital Statistics data coded with ASGCs is not recommended at a geographic resolution more detailed than the RHA level. ASGCs have been aggregated to subregional, municipal, electoral districts and other similar boundaries without success since the locational discrepancy between the case data (reported at the SGC level) and population data (postal code or DA) is too great at this level.

III. Alberta Health and Wellness Calculation Procedures

Calculating Populations

Alberta Health and Wellness relies on the Alberta Stakeholder Registry File as a basis for determining populations. This file contains the Personal Health Number (PHN)³, name, birth date, sex, and mailing address for every individual in the province of Alberta who is entitled to basic medical services under the Alberta Health Care Insurance Plan (AHCIP)

This includes the following groups:

- residents of Alberta;
- residents of Alberta living currently living temporarily elsewhere (such as students attending an educational institute outside of Alberta);
- former residents of Alberta during the first three months after they move to another Canadian province;
- persons who have 'opted out' of AHCIP;

³ Personal Health Number (PHN) is a unique number assigned to every member of the Alberta population who participates in the Health Care System. These numbers are designed to try to eliminate duplicate IDs for participants. If someone leaves the province and returns several years later, it is expected that the same PHN will be re-activated.

- persons from another country who are working or studying in Alberta on valid visas (For such persons, proof of their intent to stay in Alberta for at least 12 months must be submitted and registration is immediate);
- dependents of members of the RCMP and Armed Forces.

Not included are:

- members of the Armed Forces and RCMP;
- inmates at federal penitentiaries;
- new residents of Alberta (who have come from other provinces) during their first three months in Alberta.

In order to remain eligible for health services and keep a PHN active, all individuals (or their employers if it is part of the benefits that they receive as employees) are required to make premium payments to the Alberta Health Insurance Plan and complete and return change-of-address forms. These two requirements assure that the Registration file is regularly updated.

As noted, there may be up to a 90-day delay before an individual appears (or disappears) in the file if they have recently moved to (from) Alberta from (to) another province. People will often be registered but not eligible for coverage for those three months. Hence, if this three-month waiting time were in effect on the day for which a population count is performed, these people would not be included.

The postal code in the address⁴ is used as a basis for locating persons. When the postal code is not listed or is entered incorrectly, the community name is used to derive postal code and the population is assigned to the central postal code for that community. Ideally, address information would be used to correct the postal code, but this is not currently done.

A standard set of population files is generated quarterly and contains records for all eligible Alberta residents. Each record contains PHN, age, sex, postal code, and RHA (as derived from the Alberta Treasury PCTF) as well as certain other fields for specialized purposes. The calculation has historically occurred 6 months after the date for which the population is valid to ensure that all data have been entered into the respective databases. For population based funding projects, a four-month lag was generally used, and starting in 2000 this 4-month lag was used for all population files retroactive to 1995.

These population files are further aggregated to provide summarized data that is then used by individual branches within Alberta Health and Wellness. Because date of birth and sex are available, these summaries can be generated by age and sex, an important consideration for epidemiological calculation.

Some branches use only the fiscal year end populations (March 31). Health Surveillance typically uses the mid-year population (June 30) in accord with the established standard

⁴ This address is the mailing address, which may or may not also be the residential address.

in epidemiology for calculating yearly rates. Branches within Alberta Health and Wellness may also request a population at any time. The most common requests are associated with the calculation of funding models.

Taken together, and with appropriate caveats regarding sources of errors in coverage, this means that Alberta Health and Wellness has been able to create and maintain a postal code-based population file, something unique in Canada. It allows the calculation of populations at the Postal Code (with EA centroid) level by age and sex. If health event data is available at the postal code level, then rates and other calculations can be performed at that level⁵ at any time for any time period. Small populations make these rates unstable for a large proportion of the postal codes, however, though potential exists for the creation of intermediate geographic units between the Postal Code (with EA centroid) level and the RHA level⁶.

By contrast, health measure calculations in other provinces and by federal agencies such as Statistics Canada must use established Statistics Canada populations (generally EA census data or CSD projections for inter-census periods) as denominators. This restricts the geographic units to be employed, the time periods for calculations, and introduces potential inaccuracy and a potential for a numerator-denominator bias in rate calculations.

Counting birth and death cases

Birth and death data is collected and managed by Alberta Vital statistics. As previously described, each record is assigned an ASGC. Assigning a case to an RHA requires the ASGC-RHA cross-reference file generated by Alberta Government Services.

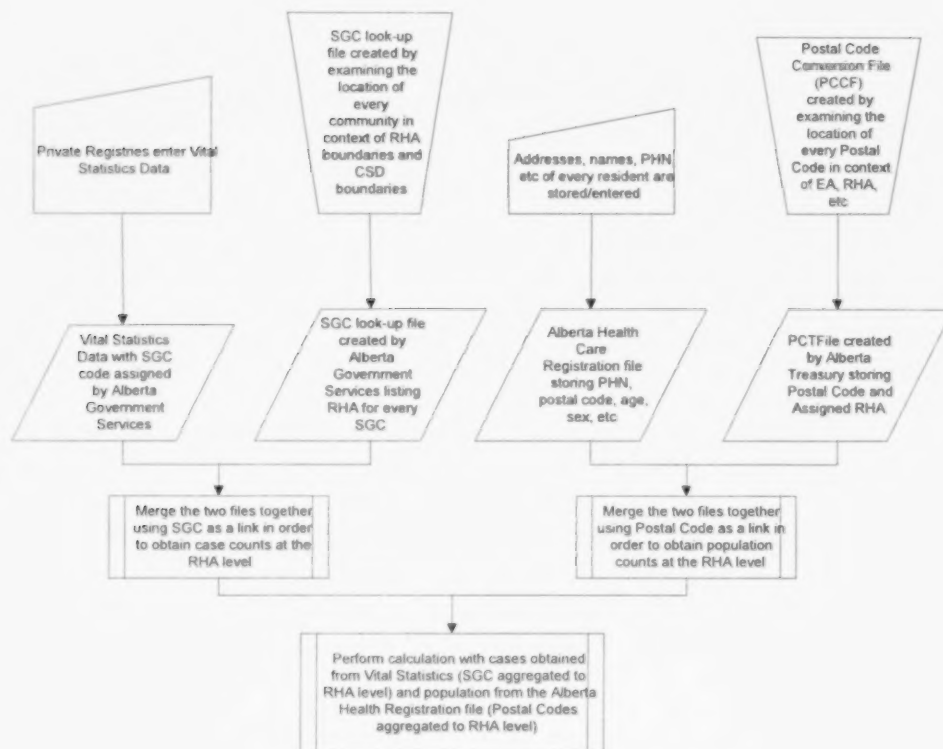
It should be noted that the Vital Statistics records do not contain as a required field the Alberta Health and Wellness PHN that would allow linkage to the Stakeholder Registry and in theory allow geographic assignment to proceed in the same fashion as described above for populations. Similarly, while Vital Statistics records do allow the collection of Postal Code, it is not a required field, and quality control procedures have not been established for it.

Figure 4 summarizes the full process for calculating birth or death rates.

⁵ In Alberta, Physician Claim, Morbidity, and Ambulatory Care databases indexed by PHN are available as a source of health event data for rate calculation. The link between the PHN and the Registry guarantees that cases in the numerator and populations in the denominator will be consistently assigned.

⁶ A project is currently in Health Surveillance underway to more fully examine this possibility.

Figure 5: Procedures used by Alberta Health and Wellness to calculate rates with data obtained from Alberta Vital Statistics



Calculating morbidity (hospitalization) cases

Data on hospitalization is obtained via the hospitalization database from the Canadian Institute for Health Information (CIHI). This file records the PHN of the person treated, the facility ID, date of admission, date of release, the codes for the treatment provided, etc. Alberta provides CIHI an RHA to postal code assignment list that they use to allocate RHA membership to every record. This method ensures that the assignment of all hospitalization records is consistent between CIHI and AH&W. By linking PHN with the registration database, it is possible to obtain the address of every individual who received treatment and thus their postal code. The PCTF file can then be used to determine the assignment to other boundaries (other than RHA) for every person who received treatment at a hospital.

Calculating claims cases

Case data can be obtained from the Physician Claims Database. This database lists every billable service provided by every physician, along with the patient's PHN. The PHN is used to link to the registration file, to obtain the patient's postal code. This is then assigned to an RHA using the PCTF. The rate calculation is performed by examining the total number of cases in the context of the relevant populations.

Comments on calculating rates

Rate calculations are reliable as long as the numerator and denominator are obtained in a consistent manner. For example, for claims data, the doctor submits a list of services provided to every PHN for a given period of time. These services are summarized using the registration file and the Alberta Treasury PCTF in order to determine counts of these health events. The population data is also generated using the registration file and the Alberta Treasury PCTF, therefore there is a consistent methodology for the top and bottom of the equation. If a few individuals in the populations have been incorrectly assigned to the wrong postal code through this process, at least the health events will be assigned in the same manner as the population was.

There remain certain challenges, such as populations and cases assigned to retired postal codes. Airdrie, for example, used the postal code "T0M0B0" until it was retired and replaced with a large number of postal codes after the community achieved city status. Changes in address were not made in several situations, therefore an inactive postal code was assigned cases and population (T0J0J0—Brooks lists a population of 6 despite the fact that the postal code has been retired for over a decade and a half). When the data is aggregated to a higher level such as a Census Sub-Division (CSD), Census Division (CD), or Regional Health Authority (RHA) the cases and population of active and retired postal codes are pooled together which eliminates the problem.

IV. Statistics Canada Calculation Procedures

Although many data sources are shared between Alberta Health and Wellness and Statistics Canada, Statistics Canada methodologies, especially for calculating populations, are sufficiently different from those employed by Alberta Health and Wellness that discrepancies can easily arise.

Calculating Populations

Statistics Canada generates population data using the population census. A full census is performed by Statistics Canada every five years. The basic census measures are collected with the short form of the census, while 20% of the population is required to fill the long version of the census, which asks questions about income, employment, occupation, language, etc. All population data is obtained from the short form resulting in close to 100% participation. The EA (DA) is the smallest geographic area for which demographic statistics are tabulated and released. Socio-economic data may be suppressed when the population for an EA is small to prevent the identification of an individual. Census data are usually not available for at least a year after the census was performed. A full description of the process, forms, fields, etc can be obtained from the Statistics Canada web site.

The population data collected at the EA level can be aggregated to the RHA level using the EA to RHA lookup file described earlier. The assignment of population by Statistics Canada is different than that used by AHW. The most obvious difference between the two methods is that the boundaries of the original sources (Postal Code for AHW, EA for Statistics Canada) vary and therefore, after amalgamation, discrepancies may be found. Further, Alberta tabulates registry population based on the mailing address (i.e. Postal Code), while the census tabulates population based on the individual's physical location. AHW aggregates postal code population to the RHA level using the Alberta Treasury PCTF, while Statistics Canada can use the EA to RHA lookup file to aggregate EA populations. A large number of geographic units (in both cases) help in reducing the overall impact of these differences.

As time from the census increases, however, populations will change, and Statistics Canada populations are generally estimates provided by projection models produced annually by the Demography Division. Such models attempt to account for births, deaths, and intraprovincial, interprovincial, and international migration. These calculations are made at the Census Division level, however, which do not map directly onto RHA units.

For the Compact Disc product entitled "Health Regions 2000" (Statistics Canada, 2000), demographic estimates for RHAs were derived using the Census Division population projections and the following steps:

1. The 1996 census population by Census Divisions (EAs or CSDs in the case of split Census Divisions) was adjusted for late changes in geography and for partially enumerated Indian Reserves to ensure that EA populations summed to the Census Division populations.
2. In cases where RHAs split Census Divisions, the proportion of population in each Census Division split was used to derive 'conversion factors' in order to allocate the Census Division level projection estimate to the appropriate health region.

Calculating birth and death cases

Statistics Canada receives Vital Statistics data from Alberta Government Services, just as AHW does. The AHW file contains ASGC assignments for every record as well as address information. While AHW uses an ASGC to RHA lookup file to assign cases to RHA, the method used by Statistics Canada is different.

The records Statistics Canada receives contain full address information, including place of residence and postal code for the mother (for births) or the deceased (for deaths). The addresses are verified using several lookup files (e.g. community name vs Postal Code, street names, active postal codes, etc). Where the postal codes are not present, they are generated using the full street address information. If an address is verified through this data quality process, it is assigned to an EA and then RHA using PCCF+ (discussed previously).

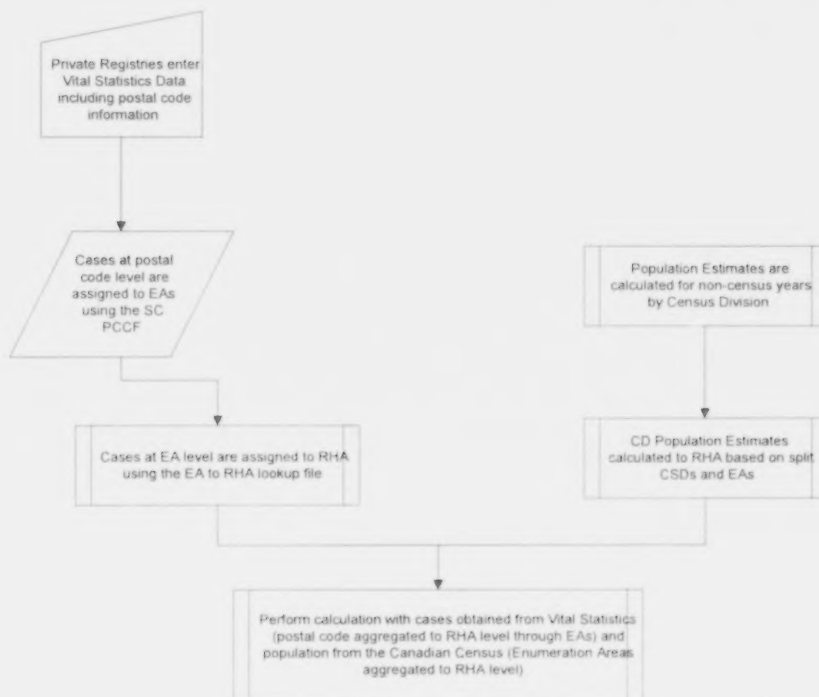
The records received by Statistics Canada also contain a CSD value inserted by Alberta Government Services using a lookup file. This CSD value is not used most of the time, but it serves as a data quality check. If the EA assigned by the process described above is within the CSD listed for the record, then it is assumed to be correct.

Statistics Canada has recently processed the Alberta Vital Statistics data for 1995 using this process. They have now fully processed records from 1995 forward.

There were measurement discrepancies between Statistics Canada and AHW calculations for Alberta 1995 death data. Upon examination, it was discovered that Statistics Canada had processed the 1995 data using a CSD-RHA assignment provided by Alberta Government Services (as the address data were not yet available). However, a change in RHA boundary in 1996 that had split a CSD between two RHAs was not properly captured in the ASGC to CSD lookup file used by Alberta Government Services. As a result cases were erroneously assigned to a single RHA. It was this problem that was responsible for the largest portion of life expectancy discrepancies between AHW and SC noted earlier. This discrepancy did not appear in calculations for later years (where the address assignment technique described above had been used).

Figure 5 summarizes this process:

Figure 6: Procedures used by Statistics Canada to calculate rates with data obtained from Alberta Vital Statistics



Calculating morbidity (hospitalization) cases

Statistics Canada had the responsibility of calculating hospital morbidity/surgical procedures rates. The responsibility for this task was assumed by the Canadian Institute for Health Information (CIHI) in the 1994-1995 data year. All analysis and assignment of records are now the responsibility of CIHI. Information about the methods used by CIHI appear in "Other Agencies" in this document.

In the Health Indicators report (December 2000), Statistics Canada generated rates at the health region level for all of Canada. The Alberta populations were first generated using Statistics Canada data, but the final document used population data generated by AH&W, based on the stakeholder's database. The AH&W population growth figures at the RHA level were applied to Statistics Canada's Demography Division's estimates for 1996 to account for the annual changes in population.

Calculating claims cases

Calculations using claims data cannot be performed by outside agencies unless they have signed a special agreement with Alberta Health and Wellness. The data is extracted for each specific agreement and included in this file are the relevant cases, time period, etc. The recipient receives postal code information for each record, if needed. The postal code information can then be assigned to EAs and then RHAs using PCCF+. The latest release of PCCF+ also provides DA assignments for every postal code.

V. Analysis of differences between calculations

Methodology

Differences in the methodology used by Statistics Canada and Alberta Health and Wellness have been described in detail above. Table 2 provides a summary.

Table 2: Comparison between Alberta Health and Wellness and Statistics Canada Methodologies

	AHW	Statistics Canada
Populations	Source: AHW Registration file Conversion: Alberta Treasury PCTF	Source: Census and projections; Conversion: EA to RHA file.
Births/Deaths	Source: Alberta Vital Statistics; Conversion: ASGC to RHA file	Source: National Vital Statistics Conversion: Postal code to EA to RHA using PCCF and PCCF+
Morbidity	Source: CIHI Morbidity Conversion: None: Alberta Treasury PCTF RHA assignment is part of the file.	Source: CIHI Morbidity Conversion: None: Alberta Treasury PCTF RHA assignment is part of the file. Morbidity rates are calculated by CIHI, not Statistics Canada.
Claims	Source: AHW Physician Claims Conversion: Link to AHW registration file and use Alberta Treasury PCTF	By special agreement only
Notes	Postal code to RHA discrepancies. Postal code to ASGC discrepancies.	EA to RHA discrepancies. CSD to ASGC discrepancies. Postal code to EA probabilistic model for rural postal codes.

Population Differences

Table 3 displays differences in population figures calculated by AHW and SC methods as described above for 1995.

Table 3: CD-based Population Estimates (Statistics Canada) and Postal Code (AHW) based populations

Health Regions Population, 1995			
Name	CD-based Population Estimates (Statistics Canada)	Stakeholder Registry Population (AHW)	% difference
Chinook RHA	143439	143778	-0.24
Palliser RHA	85761	84312	1.69
Headwaters RHA	69274	65315	5.71
Calgary RHA	825149	816238	1.08
RHA #5	51862	50607	2.42
David Thompson RHA	178263	175534	1.53
East Central RHA	103744	102017	1.66
Westview RHA	88505	84024	5.06
Crossroads RHA	38736	38534	0.52
Capital Health Authority	777895	785319	-0.95
Aspen RHA	87503	80517	7.98
Lakeland RHA	107870	105341	2.34
Mistahia RHA	84038	82515	1.81
Peace RHA	20718	20094	3.01
Keeweenok RHA	22984	24667	-7.37
Northern Lights RHA	36930	36442	1.32
Northwestern RHA	17214	17119	0.55

Disagreements of similar size are also present in other years. These discrepancies are serious enough to merit further detailed attention.

The most important possibilities for understanding these discrepancies are differences in currency, differences in geographic data source, and differences in coverage. These are separately discussed below:

Currency

Census populations are only accurate during the collection period, and estimates (projections) are needed for the period between surveys. Such estimates are less precise as the number of years after the census increases. Furthermore, a complex allocation

process must be used to estimate RHA populations since projections are generally available only for Census Divisions.

On the other hand the AHW registration file reflects current mailing address continuously. Each health care subscriber is required to register a change of address when a change occurs. This may not occur in all cases. For example, some retired postal codes report populations long after the postal code has been changed to a new code. Since mail continues to arrive despite outdated postal codes, there is little perceived need to send a change of address form when a change of residence has not occurred.

The registration file populations have been supplied to, and used by, the Demography Division to improve the accuracy of Statistics Canada population projections.

Mailing vs residential addresses

The census uses physical address in order to calculate populations for any specific EA, CSD, etc.

The registration file estimates populations based on mailing address. There are instances when the mailing and physical addresses are not the same. Some people may receive their mail at their place of work, or elsewhere. Some rural residents may list what appears to be an urban mailing address (Box 1191) while their residence is in the surrounding rural area. There are instances where some people relocate for part of the year for temporary work, but return to their listed mailing address on a regular basis. These differences may cause discrepancies in population estimates.

RCMP, military populations, and correctional institutions populations

Members of the armed forces and of the Royal Canadian Mounted Police are properly recorded in the population census, but it is not always possible to account for these populations using the Health Registration file. These populations need not deal with Alberta Health as individuals, and coverage is provided through the appropriate institution. The mailing address for all these individuals is the main central mailing address of the institution and therefore their residential status cannot be ascertained. The largest discrepancies may be observed where these populations are largest: Edmonton, Cold Lake, Suffield, etc.

Reserve populations

Estimates of First Nation's populations are always a challenge. The census does not properly reflect true populations because some bands have historically opted not to participate in the census. Furthermore, some census takers have only been allowed to contact the chief of a band and not individual band members. Further errors appear as the chief answers the questions for the entire band. Each band has its own population estimates that are used to determine funding levels from the federal government, but these are also questionable. These estimates should improve in the next census because

bands will not be able to receive certain federal funding if they do not participate in the census. For the 1996 census, 9 reserves refused to provide information for an estimated total impact of 11,287 people. The census CD population estimates add this 'missing' population into the appropriate regions.

First Nation populations can also be estimated from the AHW Registration file. Health Canada is responsible for the health and health care of Treaty Status First Nation's members, a responsibility partially discharged in the past by paying the AHCIP premium. As a result, Treaty Status members are flagged. This allows for the calculation of Treaty Status residents by postal code and for estimates of the populations. There are instances where a single band may use several postal codes, especially the Blood reserve, but these numbers are usually aggregated to the RHA level, so the impact should be minimal. Non-treaty native populations cannot currently be identified.

Non-inclusion

Both the Census and the AHCIP registration processes are required under law. There are, however, situations in which it might reasonably be suspected that individuals have not been registered under one, the other, or both. These include homeless populations, and seasonal transient and migrant individuals (perhaps more likely captured by Census). It is also possible that young healthy individuals resident in Alberta may not have been registered with the AHCIP (and have escaped the attention of the system since they have not required medical attention). Historically, census populations for the age group 20-34 have been slightly larger than in the AHW populations.

Table 4 summarizes the major potential sources of population differences.

Table 4: Potential Sources of Population Differences

	AHW	Statistics Canada
Currency	Weekly, although there may be up to a 90 day delay.	5 years with yearly estimates at CD level.
Address	Mailing address	Physical address
RCMP, Military populations	Not included	Census (every member listed)
Reserve Populations	Separate Treaty PHN available at postal code level.	Census data at Reserve level (no data available for some reserves)
Non-inclusion	Young healthy individuals (?) Homeless (?) Transient and Seasonal (?)	Homeless (?)

Vital Event Count Differences

There are two major sources of discrepancy possible for vital event counts between Alberta Health and Wellness and Statistics Canada. The first, relevant primarily to deaths, relates to the fact that the National Vital Statistics registry assembles information

from all provinces, while the Alberta Vital Statistics registry assembles information only on events that occur within Alberta. The result is that Statistics Canada can allocate deaths to Alberta regions *when these deaths occurred outside the province*. For the years 1995, 1996, and 1997, these deaths to Alberta residents numbered 231, 202 and 227 respectively. Under agreements with Statistics Canada and the provinces, records of these deaths should soon become available for inclusion in Alberta Health and Wellness calculations.

A second major source of discrepancy might relate to the differences between the allocation processes using ASGC in Alberta and the postal code to EA to RHA process used by Statistics Canada. This possibility was examined by conducting an empirical study on Alberta data.

The Vital Statistics birth data for 1995 were used. The number of births was aggregated to the RHA level using postal codes and the PCTF. For births where the postal code was missing, the community name was used to derive it. This was compared to an aggregation using the ASGC codes and the ASGC to RHA lookup file. The instances where the same birth took place in different RHAs, when aggregated, were investigated to determine the nature of the inconsistency.

Some communities have been assigned in error in the ASGC to RHA file (2001 boundaries). The most obvious error was with Calgary 06167, which is assigned to RHA 01 instead of RHA 04 (this error has since been corrected). Most of the remaining issues are associated with communities that have a rural and an urban component. These communities are near the RHA border and the rural component has been assigned to a neighbouring RHA. These discrepancies only increase when the boundaries are smaller, therefore it is not recommended that any data collected at the ASGC level be aggregated to any boundaries smaller than RHAs.

Table 5: ASGC to RHA discrepancies in the urban-rural division

Place Name	Discrepancies
Bon Accord	Urban 11/Rural 12
Calgary (06167)	In 01, should be 04
Cereal	Urban 05/Rural 01
Delburne	Urban 06/Rural 07
Gibbons	Urban 12/Rural 11
Granum	Urban 01/Rural 03
Innisfree	Urban 07/Rural 12
Milo	Urban 03/Rural 02
Minburn	Urban 07/Rural 12
Slave Lake	Urban 15/Rural 11
St. Albert	Urban 10/Rural 11
Three Hills	Urban 06/Rural 05
Vegreville	Urban 12/Rural 07

Typographical errors were found in community names for some communities⁷. While these communities are properly assigned to the RHA in the lookup file, the misspelled community names were assigned the wrong ASGC⁸. There are also cases where the name of the community and the postal code do not match up. These can be attributed to typographical errors in filling out the form, or in keypunching the records (Is that should be 4s, M vs H, and in some cases simply the wrong postal code).

Overall, there were a total of 600 discrepancies (out of a total of 38,531 records) in assigning RHAs based on postal codes compared to assignment based on ASGCs. Approximately half of the records (308) are a result of ASGC assignment errors or ASGC lookup file errors. The remaining discrepancies were the result of incorrect postal codes associated with the corresponding communities, most of which could be corrected by using community name⁹.

The results of this analysis indicate that there are very few discrepancies when Vital Statistics data are geocoded at the postal code level versus geocoding at the ASGC level. Given that many of these discrepancies can be corrected, it should generally make little difference if data is analyzed at the postal code level or at the ASGC level.

VI. Potential for numerator-denominator bias

Numerator-denominator bias may occur in the calculation of a rate or proportion when the populations included in the denominator do not accurately reflect the population at risk for the cases included in the numerator¹⁰. This includes situations in which the population at risk is undercounted as well as situations in which the denominator erroneously includes individuals not subjected to the counting processes for the numerator (at the risk of missing cases).

This is most likely to occur if the geographic units used to constrain the counting process in the numerator are not the same as those used for the denominator. It may also occur if the time periods are different for numerator and denominator.

Statistics Canada calculations include the potential for numerator-denominator bias for both of these reasons. First, the geocoding of cases relies upon the postal code, while the geocoding of denominators relies upon Census derived populations directly available at the Enumeration Area. As previously noted, these geographies do not overlap exactly. Statistics Canada attempts to overcome this potential source of bias with PCCF+ by using

⁷ These included: Athabasca, Bon Accord, Camrose, Cluny, Dixonville, Erskine, Ft Vermilion, Fox Lake, Frog Lake, Gibbons, Grimshaw, Hay Lakes, High Prairie, Hobbema, Linden, Morinville, Morrin, Olds, Rocky Mountain House, St Albert, Slave Lake, Strathmore, Sundre.

⁸ A full list of all communities and miss-assignments can be provided on request.

⁹ There are situations where the postal address contains a community name but where the actual location is within a different RHA than the community. This happens especially in the Fort Saskatchewan/Strathcona County areas. In such allocations, the RHA is correct but the community name, unfortunately, does not reflect the same RHA.

¹⁰ More properly, bias refers to the situation in which errors of this type are introduced in repeated calculations.

complex probabilistic assignment procedures for cases (geocoded by Postal Code) to EAs (DAs) whenever there are multiple EAs (DAs) served by a postal code.

Second, Census populations are collected at a single point in time. Using them as denominators for cases enumerated at a different time potentially introduces bias. Statistics Canada attempts to overcome this potential by producing population projections between census periods, but generally only for Census Divisions. The accuracy of the resulting rates depend not only upon the accuracy of the projections, but also in the adjustment procedures required because CDs do not overlap other geographic units such as RHAs.

These potential sources of bias (with the exception discussed below) do not occur in Alberta Health and Wellness procedures. First, the postal code to EA assignments are univocal, and the postal code assignment for numerator cases is accomplished through linkage (via the PHN) with the records from which the denominator populations are derived. Second, denominator populations are available at any time for any time period and can therefore be adjusted to time differences introduced in the numerator counting processes¹¹.

The exception involves the counting of Vital Events in which the geocoding for numerator cases is different for the numerator cases and denominator populations¹².

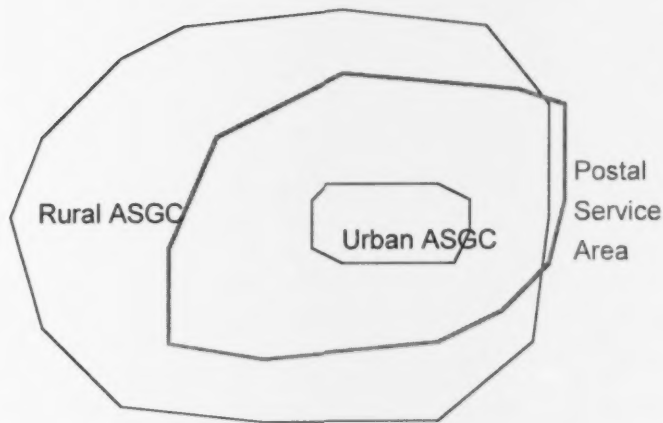
Remaining Potential for bias

Figure 6 illustrates a difficulty that might arise in working with data collected at the ASGC level along with data collected at the postal code level in any small area analysis. The postal code service area includes all the urban ASGC region, but only part of the surrounding rural ASGC. The problem could be magnified if the geographic coordinates of the postal code are derived from the EA as they are in the PCTF.

¹¹ There does remain a potential for bias in the EA to RHA aggregation since the overlap is not perfect, but this source of bias also exists in Statistics Canada calculations.

¹² As discussed previously, however, empirical work suggest that the impact will be minimal in aggregating to the RHA level.

Figure 7: Postal Code Area vs. ASGC Area



Appendix 2 provides a case study that illustrates other problems that an analyst might encounter in performing small area analysis.

VII. Other Agencies

Alberta Cancer Board

The Alberta Cancer Board (ACB) maintains a cancer registry for the Province of Alberta. It is considered the best source of cancer data in the Province of Alberta. ACB regularly calculates epidemiological quantities related to cancer surveillance.

Incidence

The Cancer Registry presents an accurate and up-to-date view of cancer. All laboratories are required to send the ACB a copy of each pathology report with a diagnosis of cancer. These are placed in registry file that stores historical information for every member of the population that has been diagnosed with cancer. It is thus possible to determine if a case is new, a re-occurrence in an existing cancer site, or a re-occurrence in a different cancer site.

Population

Population data at the provincial level is used to compare measures in Alberta with other provinces. For these statistics, the Statistics Canada 1991 and 1996 census data are used. For non-censal years, Statistics Canada projections are used. For calculations at the RHA level, the Alberta Health and Wellness populations have been used. These are derived from the stakeholder mailing address and aggregated to the RHA level. No changes are

made to this file, so these are the same populations that are used for AHW reports and analysis.

Mortality

The mortality data used by the ACB is obtained from Vital Statistics. This data source is identical to that used by AHW.

Geographic Assignment

Case data is obtained at the postal code level and these are aggregated to the RHA level using the Statistics Canada Postal Code Conversion File using the Single Link Indicator to assign all postal codes to a single, most representative, Enumeration Area. These are then assigned to the RHA level.

Mortality data is normally assigned to an Alberta Standard Geographic Code (ASGC). These do not correspond very well to postal codes for sub-regional analysis. However all documents published by the ACB to date contain RHA level where the potential discrepancies are fewer.

In summary, the population data at the RHA level is consistent with AHW. The mortality data is also consistent as they are derived from the same source and use the same geographic locators. The ACB does have an advantage over AHW for mortality data in that all deaths can be cross-referenced against the cancer registry to investigate any possible errors in coding the cause of death.

Incidence data is also more detailed at the ACB, however the assignment of cases to the RHA level differs from the methodology used by AHW. The ACB uses the Statistics Canada Postal Code File, while Alberta used the Treasury Postal Code File. The differences in these two lookup files can create problems for small area analysis, however are unlikely to create major discrepancies at the RHA level.

The ACB does generate special analysis at the postal code level on request. If the same data were to be used but the AHW methods used, then some discrepancies would be expected as a result of different postal code assignments. Furthermore, any analysis of mortality data at the postal code level would have to be based on postal codes obtained from the registry, not the ASGC assignment as these are not congruous with postal codes.

Alberta Health and Wellness, Finance and Corporate Services

Alberta Health and Wellness, Finance and Corporate Services base a number of decisions at the community level based on postal code data. These calculations are usually associated with funding issues, import-export figures among RHAs, and similar issues and are all at the RHA level. Some of the models have been written to work within spreadsheets, however MS Excel is not able to store all of the postal codes of the

province in a single worksheet. AH&W Finance summarized all of the postal code data to the community level which results in less than 1,000 unique geographic entities using the Alberta Treasury postal code conversion file (PCTF). Each postal code is associated with a community name, therefore this community name is used to summarize all of the data for the community. For example, Brooks summarizes all of the TIR postal codes as well as the 6 people who still report their postal code as T0J0J0 (also Brooks). Since the majority of the postal codes appear in urban areas, this allows for data summarization without any loss of geographic correspondence as long as the data is aggregated to the RHA level. In those instances where the postal code for a community falls within more than a single RHA, the community name appears as many times as required. For example, with the 2001 RHA boundaries, the City of Fort Saskatchewan was part of region 12 despite being physically separated from the rest of the RHA. As a result, two thirds of the postal codes were in RHA 12 and the other third were in neighboring region 10 (Capital). The community name of Ft Saskatchewan was split into two names: Ft Saskatchewan(12) and Ft Saskatchewan(10). The appropriate postal codes were assigned to the corresponding portion of the community. The same approach was used in other areas where the postal code's community appeared in more than a single RHA.XI.

Canadian Institute for Health Information (CIHI)

The Canadian Institute for Health Information (CIHI) has the responsibility for all morbidity/surgical procedures data collection and management starting in the 1994-1995 data year. This file records an identifier for the person treated, his or her address (including postal code), the facility ID, date of admission, date of release, codes for the treatment provided, etc. Alberta provides CIHI an RHA to postal code assignment list that they use to allocate RHA membership to every record for their Discharge Abstract Database (DAD). This method ensures that the RHA assignment of all hospitalization records is consistent between CIHI and AH&W. Records without a complete postal code or with a postal code that cannot be matched to Alberta's PCTF are excluded from the region assignment and rate calculation.

In the Health Indicators report (December 2000), Statistics Canada generated rates at the health region level for all of Canada. The Alberta populations were first generated using Statistics Canada data, but the final document used population data generated by AH&W, based on the stakeholder's database. The AH&W population growth figures at the RHA level were applied to Statistics Canada's Demography Division's estimates for 1996 to account for the annual changes in population.

For all other work, CIHI usually calculates rates at a provincial level and therefore uses Statistics Canada data to ensure nation-wide consistency. Statistics Canada adjusts the provincial population estimates using the AH&W population information for non-census years. Discrepancies are bound to continue for reasons already outlined in this document, such as data collection date, RCMP population, military population, First Nations population, etc.

VIII. Conclusions and Recommendations

This report has outlined the processes used to generate health measures by AHW, Statistics Canada, and other agencies. Although these methods do differ quite extensively, the limited number and size of discrepancies should be heartening to analysts.

Alberta is in the fortunate and unusual position of being able to use the Stakeholder Registration database as a source of ongoing population by age, sex, and full address. Any problems introduced by less-than-optimum postal conversion or translation can usually be applied equally to both numerator and denominator or rates and proportions to be calculated, so that a major source of potential bias can be avoided. In other provinces, the Stakeholder Registry file is not a good reflection of the resident information or addresses because it is not usually up-to-date. Many of the comments made about the suitability of a particular file or methodology are, therefore, only applicable to Alberta and extensive research and analysis are needed before proceeding with these techniques in other jurisdictions.

Ideally, AHW would be able to treat the Vital Statistics counts at the postal code level. However, such a change in procedure must await changes in the manner in which Vital Statistics data is collected in Alberta. It will be necessary either to allow the collection of the individual's PHN to allow linkage of his or her record to the Registration file, or to require a postal code on Vital Statistics address fields and subject it to verification procedures. Fortunately, existing procedures appear not to introduce large error at least when aggregated to the RHA level.

Perhaps more importantly, AHW must take steps to include the records from Albertans who die or give birth in other provinces in health measure calculations.

The potential also exists in Alberta to create levels of geography intermediate between the Postal Code/EA level and the RHA level, and to examine counts and rates at these levels. Unique advantages for all levels of governance might be obtained if these units were to form a nested hierarchy from Postal Code/EA to RHA. Of course, care would need to be taken to insure that the sources of potential bias addressed in this report are taken into account.

Despite the limitations imposed upon Statistics Canada population figures (i.e. that they are available at the EA (DA) level and only at a single point in time), Statistics Canada has spent great energy in the development and promotion of techniques that eliminate as much bias as possible in the calculation of health measures. These include the development of PCCF+ as a mechanism to eliminate bias when converting geocoding at the postal code level to coding at the EA level, and the development of a program for creating population projections to improve the accuracy of population data as time since a census increases.

The principle recommendation include

1. Examining the potential to assign vital event records using postal code.
2. Entering agreements with Statistics Canada and other provinces to obtain Vital Events records when these events occur in other provinces in order to include these counts in the calculation of health measures.
3. Investigating the potential for creating intermediate levels of geography for the calculation of health measures.
4. Encouraging Alberta Treasury and Statistics Canada to create a consistent Postal code file for Alberta.
5. Encouraging Statistics Canada to re-examine methods of assigning projection populations to Regional Health Authorities; and making every effort to improve the quality of these projections by incorporating information available from the Alberta Health and Wellness population series.

Glossary

AHW	Alberta Health and Wellness
ASGC	Geographic code assigned to Vital Statistics records in Alberta. <i>Not to be confused with SGC.</i>
Block Face	Smallest geographic unit in the census, for which total population and dwelling counts are available. Block faces are only available where there is Street network File coverage.
Census	Population census, performed every 5 years.
CIHI	Canadian Institute for Health Information, agency responsible for the collection and storage of all hospital discharge information in Canada.
CD	Census Division - Used to report census data at a higher level of aggregation than CSD.
CSD	Census Sub-Division -Used to report census data at a higher level of aggregation. These correspond to the county, MD, SA, ID, city, town, and reserves boundaries.
Community name	Name of the community as it appears in the address field of Vital Statistics, Health Registration, and other records.
EA	Enumeration Area - The basic unit of data collection for the population census. This corresponds to the information collected by a single census-taker. It is the smallest area for which data is compiled and released
LDU	Local Delivery Unit, the location based on the full six digit postal code (block face or EA centroid).
Mailing address	Address at which mail is received for health plan payments. It does not necessarily match the residential address.
numerator-denominator bias	a systematic error that may be introduced in the calculation of a rate or proportion when the populations included in the denominator do not accurately reflect the population at risk for the cases included in the numerator
Postal code	6 digit code that identifies a unique location for every mail delivery destination. Rural postal codes are identified with a "0" in the second digit and normally a much larger geographic region than an urban postal code.

Postal code area	The region served by a postal code, for rural postal codes it includes the urban and rural component of the postal code.
Postal code centroid:	Centroids are a unique combination of latitude and longitude coordinates that define a point location somewhere inside a polygon or region.(The point does not necessarily need to be located in the center of the polygon, despite the name- convention usually places the point near the center or in the center of the larger portion of the polygon).
Postal code file	A file that lists all postal codes, the corresponding community, geographic coordinates, and other relevant information.
PCCF	Postal Code Conversion File, refers to the postal code file created by Statistics Canada.
PCCF+	A set of SAS program written to allow the analysis of population data and census-based information at the postal code level. It is based on a population-weighted probabilistic method applied to the PCCF.
PCTF	Postal Code Translation File, refers to the postal code file created by Alberta Treasury.
Rate	Number of cases over the respective population, normally expressed in per 100,000. This includes the crude and age-sex standardized rates.
Registration file	Alberta Health maintains a registration of citizens who have a Personal Health Number (PHN). This is used for billing purposes and to generate population at the postal code level.
RHA	Regional Health Authority –The entities responsible for the provision of health services in the province of Alberta. There are 17 RHAs in the province.
SC	Statistics Canada
SGC	Statistics Canada CSDs are called SGCs at a national level, but have no direct relation to Alberta codes (ASGC). The occasional use of the name SGC to refer to the Alberta codes has caused great confusion.
SLI	Single Link Indicator, the EA or block face that best represents the postal code. This is the element selected when using a single location to represent the entire postal code (single link method).
SNF	Street Network File is a GIS-ready file which contains the single line streets, railways, and water features for large urban centres, that is municipalities with populations greater than 50,000.

Bibliography

Assessment of Respiratory Disorders in Relation to Solution Gas Flaring Activities in Alberta. Health Surveillance, Alberta Health and Wellness (1998).

Ellehoj, E.A., Schopflocher, D, et al (2004). *Using the AH&W Lookup File to Aggregate Data to Different Geographic Boundaries-Geographic Methodology Series No. 2.*

Ellehoj, E.A., Schopflocher, D, et al (2004). *Health Rate Mapping Template-Geographic Methodology Series No. 3.*

Ellehoj, E.A., Schopflocher, D, et al (2004). *Calculating Distances in Alberta-Geographic Methodology Series No. 4.*

Ellehoj, E.A., Wilson, S, et al (2005). *Definition of Rural and Urban Areas for Health Services Provision-Geographic Methodology Series (in preparation).*

Ellehoj, E.A., Schopflocher, D, et al (2005). *Geographic Activities in AH&W - Geographic Methodology Series (in preparation).*

Health Trends in Alberta, 2001. Alberta Health Surveillance, Alberta Health and Wellness. ISSN 1480-6657

Geocodes / PCCF Version 3E: User's Guide (Automated Geographic Coding Based on the Statistics Canada Postal Code Conversion Files), Russell Wilkins, Health Analysis and Modeling Group, Social and Economic Studies Division, Statistics Canada, Ottawa, (2000)

How Healthy are Canadians? A Special Issue, Health Reports Vol 11 #3. Statistics Canada (1999)

Health Regions 2000 -Boundaries, Geographic Information, and Population Estimates. Statistics Canada (2000) Catalogue No. 82F0082XCB.

Life Expectancy as a Health Status Measure, Health Surveillance, Alberta Health and Wellness (1998).

Registration Data as Population Estimates, Information Dissemination, Alberta Health and Wellness (2000).

References for Data Sources

Alberta Health and Wellness publishes many reports on health issues in Alberta
www.health.gov.ab.ca

Alberta Treasury. This department is responsible for the most comprehensive and up-to-date Postal Code Conversion File available in the Province of Alberta and is used by most provincial departments and RHAs. For information contact Todd Churney (Todd.Churney@gov.ab.ca)

Canada Post has a good online resource to obtain information about individual postal codes www.postescanada.ca

Desktop Mapping Technology Inc., creators of an enhanced version of the Postal Code Conversion File distributed by Statistics Canada: www.desktopmapping.com They also sell census data and boundaries and other related geographic information. Another company offering a similar service can be found at www.tetrad.com

Government Services. This department is responsible for entering and compiling the Vital Statistics data (gail.brese@gov.ab.ca)

Statistics Canada make available census data, census boundaries, postal code conversion file, special studies, and countless reports. www.statcan.ca

Appendix 1: Sample Files

Selected rows and columns from *ASGC to RHA Lookup file*

Name: ASGC lookup file
 Elements: ASGC Code, Location Name, RHA, CSD
 Developed By: Alberta Government Services
 Maintained by: Alberta Government Services / Alberta Treasury
 Updates: Daily
 Special Notes: New ASGC codes are created as needed, but all the necessary may not be entered until it has been processed by Alberta Treasury (usually once every 5 years)

SGC_CODE	LOCATION	RHA_98
13350	ABEE	12
04010	ACADIA VALLEY	02
05443	ACME	05
05401	ACME (RURAL)	05
13270	ACREAGE	11
15940	ADAM'S LANDING	17
01070	ADEN	02
05230	AERIAL	05
03601	AETNA	01
15239	AGGIE	15
06213	AIRDRIE	04
06130	AIRDRIE (RURAL)	04
06165	ALBERT PARK	04
13123	ALBERTA BEACH	11
13001	ALBERTA BEACH (RURAL)	11
11580	ALCOMDALE	11
10350	ALCURVE	07
11001	ALDER FLATS	09

Postal Code Conversion File (Alberta Treasury)

Name: Alberta Treasury Postal Code Conversion File
 Elements: Postal Code, EA, CSD, Active status, RHA, Location Name, Start & end date, etc
 Developed By: Alberta Treasury
 Maintained by: Alberta Treasury
 Updates: Monthly
 Special Notes: Most current and complete file available in Alberta

PC	FEDEA	RETIR ED	MUNICIPALI	POPDE NSITY	LOCN AME	STA RT	END	CSD	RHA98
T0A0A0	002216	A	ABEE	R				13036	12
T0A0B0	002464	A	ARDMORE	R				12004	12
T0A0C0	002164	A	ASHMONT	R				12014	12
T0A0E0	001033	A	ATMORE	R				13044	11
T0A0G0	002454	I		L	GRCT		30-Apr-98	12004	12
T0A0H0	002007	I	CHERRY GROVE	R			30-Apr-98	12004	12
T0A0J0	002156	A	BELLIS	R				12022	12
T0A0K0	002319	A	BON ACCORD	R				11066	12
T0A0L0	002457	I	BONNYVILLE	R			30-Apr-98	12009	12
T0A0M0	001026	A	BOYLE	R				13044	11
T0A0N0	002351	A	BOYNE LAKE	R				12014	12
T0A0P0	001160	A	BREYNAT	R				13044	12
T0A0R0	001025	A	CASLAN	R				13044	12
T0A0S0	001402	I	CHARD	R			30-Apr-98	16037	16
T0A0T0	002051	A	CHERRY GROVE	R				12004	12
T0A0V0	002455	I	COLD LAKE	R			30-Apr-98	12004	12
T0A0V1	002057	I	COLD LAKE	R			30-Apr-98	12006	12
T0A0V2	002055	I	COLD LAKE	R			30-Apr-98	12006	12
T0A0V3	002056	I	COLD LAKE	R			30-Apr-98	12006	12
T0A0V4	002058	I	COLD LAKE	R			30-Apr-98	12006	12
T0A0W0	001370	I	CONKLIN	R			30-Apr-98	16037	16
T0A0X0	002315	I		R			30-Apr-98	11059	11
T0A0Y0	001021	I		R			30-Apr-98	13044	11
T0A0Z0	002263	A	EGREMONT	R				13036	12
T0A1A0	002012	A	ELK POINT	R				12016	12
T0A1B0	001032	A	ELLSCOTT	R				13044	11
T0A1C0	001403	I	CHARD	R			30-Apr-98	16037	16
T0A1E0	002105	A	FOISY	R				12014	12
T0A1G0	001425	I	FORT CHIPEWYAN	R			30-Apr-98	16828	16
T0A1H0	002451	A	FORT KENT	R				12004	12

EA (& CSD) to RHA lookup file

Name: Census to RHA lookup file
 Elements: EA, CSD, RHA
 Developed By: Statistics Canada
 Maintained by: Statistics Canada
 Updates: Every 5 years or when RHA boundaries change
 Special Notes: This file is created using a simple overlay command in a GIS and similar files have been created by AHW. These have been verified as identical to those used by Statistics Canada.

FEDEA	CSD	FEDEA POP	FEDEAD WELL	RHA
001001	11059	57		20 11
001002	11059	34		12 11
001003	13028	211		77 11
001004	13028	280		99 11
001005	13028	491		160 11
001006	13032	410		161 11
001007	13028	433		155 11
001008	13028	394		144 11
001009	13036	92		35 12
001010	13044	145		55 11
001011	13028	9		5 11
001012	13028	7		4 11
001013	13028	188		61 11
001014	13028	179		68 11
001015	13044	311		106 11
001016	13044	418		153 11
001017	13044	403		140 11
001018	13044	107		52 11
001019	13044	13		5 11
001020	13044	14		10 11
001021	13044	14		4 11
001022	13036	55		22 12
001023	13044	376		135 11
001024	13046	802		323 11
001025	13044	108		42 12
001026	13044	323		125 11

Selected records and fields from *Vital Statistics Mortality file*. All entries denoted with "xxxx" have been suppressed to ensure the confidentiality of these records.

Name: Vital Statistics Mortality File
 Elements: See below
 Developed By: Government Services
 Maintained by: Government Services and private registries
 Updates: Daily
 Special Notes:

SEX	DETH DATE	M A	PL_S GC	RES_S GC	DOB	OCCU PATI ON	TYP_ BUS	N_C AUS E	U_C AUS E	PL_R NJ Y	R R	PL_A R H	AGE YR	CA US	C H
M	xxxx	1	07022	07022	xxxx	xxxx	xxxx		514	10	07	07	19	xxxx	514 08
M	xxxx	1	11482	11182	xxxx	xxxx	xxxx	9696	9803	10	08	08	06	xxxx	980 17
F	xxxx	3	11611	11611	xxxx	xxxx	xxxx		1809	10	10	10	19	xxxx	180 02
M	xxxx	1	11611	15382	xxxx	xxxx	xxxx	8208	888	8	14	10	17	xxxx	888 17
M	xxxx	1	11611	11611	xxxx	xxxx	xxxx		4140	10	10	10	15	xxxx	414 07
F	xxxx	3	11611	11350	xxxx	xxxx	xxxx		486	10	08	10	18	xxxx	486 08
M	xxxx	2	11611	11611	xxxx	xxxx	xxxx		4148	10	10	10	13	xxxx	414 07
F	xxxx	3	11611	11611	xxxx	xxxx	xxxx		410	10	10	10	18	xxxx	410 07
M	xxxx	3	11562	11643	xxxx	xxxx	xxxx		4149	10	12	12	15	xxxx	414 07
F	xxxx	1	11611	11621	xxxx	xxxx	xxxx		4289	10	10	10	16	xxxx	428 07
M	xxxx	1	13393	13393	xxxx	xxxx	xxxx		410	10	12	12	19	xxxx	410 07
M	xxxx	2	11611	11611	xxxx	xxxx	xxxx		4409	10	10	10	19	xxxx	440 07
M	xxxx	2	11611	11611	xxxx	xxxx	xxxx		4292	10	10	10	15	xxxx	429 07
F	xxxx	2	11611	11611	xxxx	xxxx	xxxx		586	10	10	10	15	xxxx	586 10
F	xxxx	4	06161	06161	xxxx	xxxx	xxxx		1629	10	04	04	15	xxxx	162 02
F	xxxx	3	06161	06161	xxxx	xxxx	xxxx		410	10	04	04	19	xxxx	410 07
F	xxxx	3	06161	06161	xxxx	xxxx	xxxx		2500	10	04	04	18	xxxx	250 03
M	xxxx	2	06140	06322	xxxx	xxxx	xxxx	8628	8160	6	05	04	06	xxxx	816 17
M	xxxx	3	06161	06161	xxxx	xxxx	xxxx		410	10	04	04	15	xxxx	410 07
M	xxxx	2	02121	02121	xxxx	xxxx	xxxx		586	10	01	01	18	xxxx	586 10
M	xxxx	2	01061	01061	xxxx	xxxx	xxxx		4140	10	02	02	17	xxxx	414 07
M	xxxx	3	02363		xxxx	xxxx	xxxx		4140	10	98	02	14	xxxx	414 07
M	xxxx	3	01061	01061	xxxx	xxxx	xxxx		4140	10	02	02	19	xxxx	414 07
M	xxxx	2	11611		xxxx	xxxx	xxxx		430	10	97	10	09	xxxx	430 07

Appendix 2: Challenges for Small Area Geographic Analysis in Alberta

The following figures use the same geographic area to demonstrate some of the challenges in performing small area health measures analysis. Especially important to note is that discrepancies in the results obtained between various methods discussed in the current report will be magnified when geographic units smaller than RHAs are employed. This appendix does assume knowledge of the contents of the report.

Figure 8 Census Sub-Divisions with postal codes (red diamonds) Source: PCTF



Figure 7 shows CSDs (equivalent to cities, towns, counties, MDs, etc). A larger community (Vegreville) contains several postal codes. There are more than the 8 postal codes that appear in the town boundary, but these are not visible because they share a single assigned geographic location. Some of the Vegreville postal codes also provide service to the region near the town (in Minburn County). Lavoy contains a single postal code that serves the community and surrounding rural area. The community of Ranfurly

is too small to warrant CSD designation, but it does contain a postal code for the urban and rural residents of the area.

Figure 9 Enumeration areas and postal codes (red diamonds) Source: PCTF.

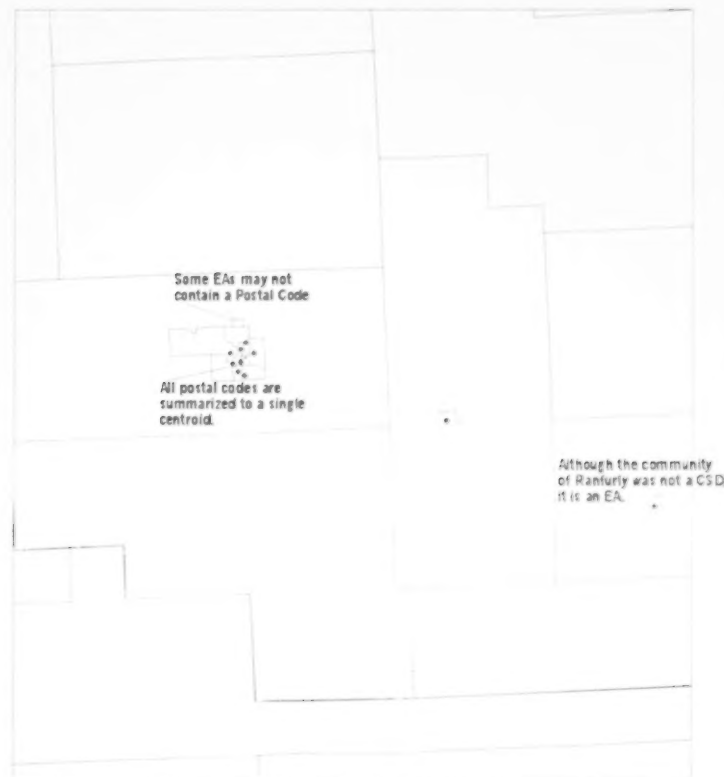


Figure 8 illustrates how the postal codes are assigned a location in the PCTF based on EAs. Each EA contains what appears to be a maximum of one postal code, but many postal codes are displayed as sharing the same geographic location. Some EAs may not contain any postal codes. (*Note: In some cases CSD and EA boundaries are identical*) The outer boundaries for Vegreville are identical, but the community is sub-divided into smaller regions. Lavoy, on the other hand is too small to sub-divide and therefore the CSD and EA boundaries are identical. There is an EA boundary surrounding Ranfurly whereas at the CSD level Ranfurly is not delineated.

Figure 10: Enumeration Areas and postal code locations (green triangles) Source: Canada Post



Figure 9 shows the same information as Figure 8, but the postal code locations are based on the Canada Post file. Lavoy's postal code appears in a consistent location, but this is not the case with postal codes in Vegreville (moved south) and Ranfurly (moved north).

Figure 11: Enumeration Area populations Source: Statistics Canada

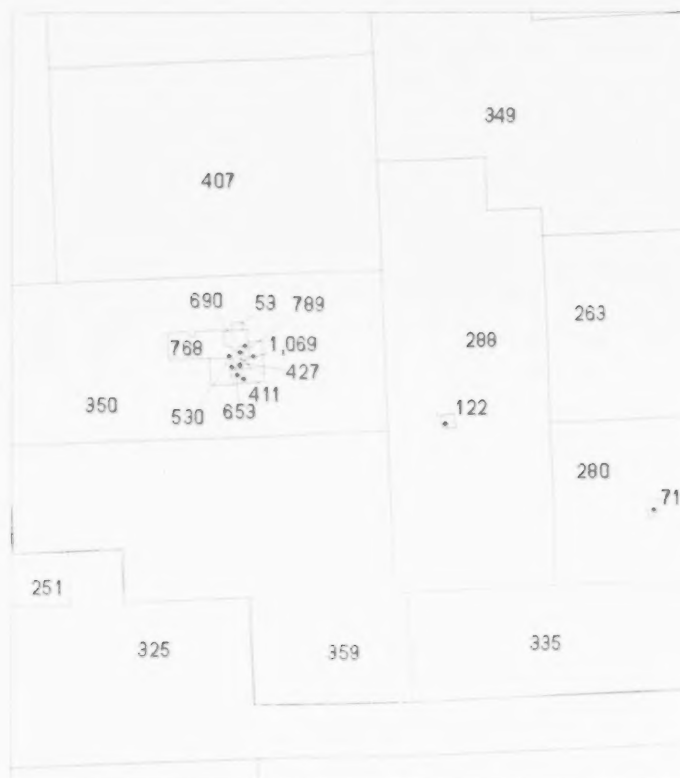
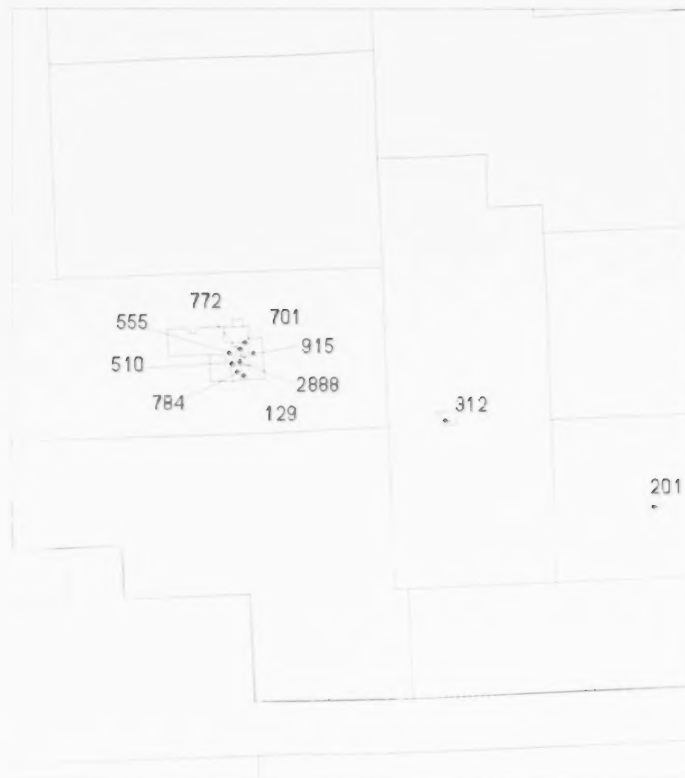


Figure 10 shows the populations for each of the EAs that were shown in the previous figures. Both rural and urban areas have populations. This figure is contrasted with Figure 11 that shows populations at the postal code level.

Figure 12: Postal code populations Source: Alberta Health and Wellness registration file



The numbers presented in Figures 10 and 11 differ dramatically. The total population presented in Figure 10 is 8,790, while the population presented in Figure 11 is only 7,767. The postal code populations show no population in the rural areas as all population is aggregated to the urban areas where the postal code office is located.

The postal code population of Lavoy is 312 while the rural and urban EA population for Lavoy is 410 (288 + 122). The discrepancy occurs because in postal code populations there is a gravitational effect associated with population size. Larger communities influence or pull more rural residents than smaller communities. In this case, many of the people located in the rural EA surrounding Lavoy use Vegreville as a post office. In the case of rural mail delivery, larger communities are responsible for the distribution of mail to the surrounding rural area. The same problem appears with Ranfurly where a population of 201 is contrasted with a population of 351 (71 + 281).

The EA populations partially reflect population for postal codes just outside the map area (Mundare, Two Hills, and Viking). If the mapping window were to be increased slightly to include those communities at the corners of the map, then the postal code populations would become a little larger than the EA populations because they would include population for a rural area that is out-of-sight.

This analysis clearly suggests that when health data is collected at the postal code level, the population used to calculate a rate should also be based on postal codes, most especially for small area analysis.

In some cases, the population differences may transcend RHA boundaries. Several RHAs that surround RHA 10 have smaller populations when using the Registration file than would be the case when using EA boundaries. Once again, these discrepancies can be attributed to the gravitational effect that larger communities exert on small communities. RHA 10 experiences this effect more than RHA 4 (Calgary). This is a result of the RHA boundaries which separates some communities from the greater Edmonton economic zone despite the great interaction between the communities. The only discrepancy near the Calgary RHA is with Headwaters RHA as a result of a boundary that divides the Stoney Indian Reserves. A similar problem exists between RHAs 9 and 6 where the boundary crosses several Indian Reserves (mostly in the Hobbema area).